z e t e r a

Zetera White Paper on

µSAN™ Using Port Addressing

An overview, compare and contrast, technical paper on using port addressing to implement the µSAN™ protocol instead of IP addressing.

This document references the µSAN™ White Paper (Version .35) an assumes the reader is familiar with the protocol.

Version 0.10

Author:
Thomas E. Ludwig
VP Engineering, Zetera Corporation

January 2003

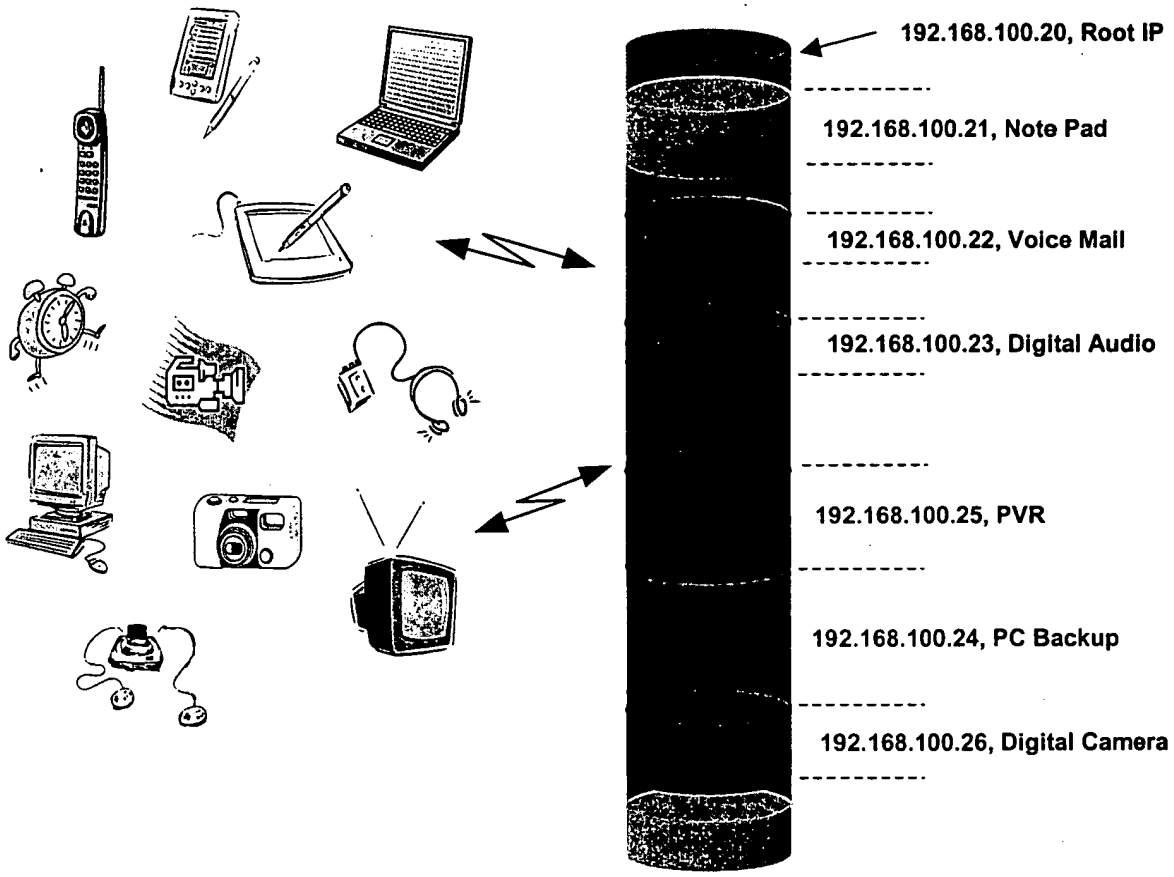
Revision History

Version 0.1	October 2002	Original Release

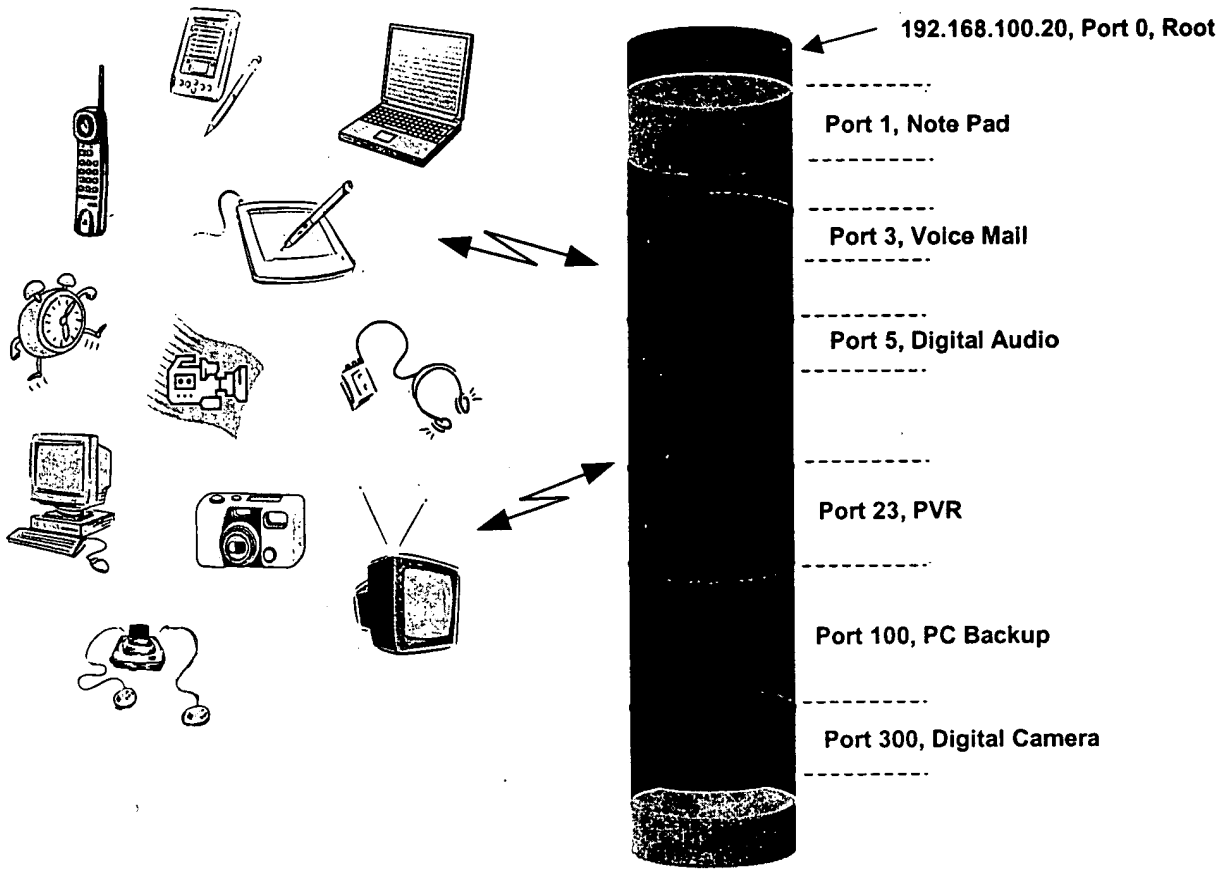
Contents

Overview

The μSAN™ White Paper identifies a storage system and protocol whereas each of the individual partitions within that storage system is addressed in the Internetwork Protocol (IP) stack by a mutually exclusive IP address as shown below. The advantages of using the IP address for partition identification lies in its exclusivity in the IP stack allowing for simple routability, masterless control and unlimited multicast associations.



This paper identifies an alternative architecture whereas instead of using IP addresses to address individual partitions, port addresses (or derivatives of port addresses are used). This type of architecture has some limited advantages and significant disadvantages compared with IP address designation. It should be understood that using the multiplexing characteristics of the Transport Layer (TCP or UDP) may also be expanded into a virtual Transport Layer in the application with similar effect, functionality and limitation. For discussion, the new network topology would look as follows:



Network Topology

In this network topology the following rules are applied:

- Each individual μSAN™ I/O uses a single unicast IP address (with a single associated multicast IP address f necessary).
    - The IP is instantiated through DHCP or Auto IP as in UPnP
    - The Root is communicated to through a fixed Port Address (show here as Port 0)
- Each μSAN™ partition is built as before, however, it is now assigned a port address instead of an IP address. These port addresses are assigned at the discretion of the μSAN™. (In a master/slave architecture, the client could designate the Port Address.)
- This architecture could work in either the TCP or UDP transport domain.
- It is assumed there will be no Port Address conflict with other applications due an isolated IP address.
- It is further perceived that a single Port Address maybe used with the port multiplex functionality performed at the application layer with similar results.
    - In this case, however, the partition address would need to be included in the Command/Data packets.
- Name resolution will not only have to pass the root IP address, but the port address as well.

Advantages

The advantages to using port addressability lies in the simplicity of the implementation. Most existing stack engines expect a single IP with ports addressing individual applications. Using a single IP means that the more complicated reentrant code will lie at the application layer. Other advantages are:

- Less IP addresses need to be assigned in the overall network.
- Only one DHCP timer needs to be maintained.

Disadvantages

The disadvantages in this architecture lie in the lack of a virtual nature of the storage, and a serious if not prohibitive nature to multicast implementation.

Virtual Nature

- Each partition is only virtual through the root IP. When each partition has it's own IP address, locality and relationship with the physical is immaterial and easily relocateable.
- To move a partition to another IP root would require port re-instantiation and would have to be different.

Name Resolution

- Typically Name Resolution protocol resolves the IP address. Here, the port must be resolved as well.

Multicast

- Multicast operations (operations to more than one virtual partition) must not only be served a multicast IP address, but a multicast port address as well. This multi-port address MUST be the same address regardless of the unicast root IP address. There is no agent in the IP protocol system to administer the multicast port address. A potential scheme for multicast port resolution is described in this document.
- Without an dedicated server and IP ratified protocol to resolve the port address assignments, (in other words, a duplication of the IPGM protocol but for ports), it is not possible to always assign port addresses that are not in conflict.

Sample Multicast Port Assignment

The following rules describe a potential scheme to assign multicast port assignment for a multicast IP/multicast port communication.

- Assume that each partition has a unicast port address assigned to it as well as a multicast port address assigned to it. The unicast port address is assigned randomly by the μSAN™. The multicast port address is assigned by the client.
- Assume that the multicast IP is achieved as described in the μSAN™ White Paper (version .35)
- Each μSAN™ that recognizes its multicast IP address would then use for partition resolution the multicast port assignment. (Note, there may be more than one of these port number matches in a μSAN™.)
- The client assigning the multicast port address to the partition would have to resolve conflict with a number already assigned. This is true both within the root IP domain as well as between μSAN™s. A method of interrogation would need to be created with in the protocol. It is possible that NO port number would be available between all of the μSAN™s.
- A method would need to be created that dissolved multicast port assignment. This would need to be done, most likely, through an outside server.